

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## (i) APPLICANT:

- (A) NAME: PLANT GENETIC SYSTEMS N.V.
- (B) STREET: Plateaustraat 22
- (C) CITY: Gent
- (E) COUNTRY: Belgium
- (F) POSTAL CODE (ZIP): 9000
- (G) TELEPHONE: 32 9 2358454
- (H) TELEFAX: 32 9 224 06 94
- (I) TELEX: 11.361 Pgsgen

(ii) TITLE OF INVENTION: New Bacillus thuringiensis strains  
and their insecticidal proteins

(iii) NUMBER OF SEQUENCES: 4

## (iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.25(EPO)

## (2) INFORMATION FOR SEQ ID NO: 1:

## (i) SEQUENCE CHARACTERISTICS:

- (A) SEQUENCE LENGTH : 19 nucleotides
- (B) SEQUENCE TYPE : nucleic acid probe
- (C) STRANDEDNESS : single stranded
- (D) TOPOLOGY : linear

(ii) MOLECULE TYPE: synthetic DNA

(ix) FEATURES: the probe is a part of the coding DNA  
strand of the cryIG gene, described by Smulevitch et al.  
(1991).

(x) PROPERTIES : this probe is used to isolate the  
bTS02618A gene from its containing strain.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

5'-TTCTGTACTATTGATTGTA-3'

## (3) INFORMATION FOR SEQ ID NO: 2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1561 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Bacillus thuringiensis
  - (B) STRAIN: BTS02618A
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (D) OTHER INFORMATION: /function= "contains the translation initiation codon of the bTS02618A gene"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

AAAAAGAAAT AGGAATAAAT ACTATCCATT TTTTCAAGAA ATATTTTTTTT ATTAGAAAGG	60
AATCTTTCTT ACACGGGAAA ATCCTAAGAT TGAGAGTAAA GATATATATA TATAAATACA	120
ATAAAGAGTT TGTCAGGATT TTTGAAAGAT ATGATATGAA CATGCACTAG ATTTATAGTA	180
TAGGAGGAAA AAGTATGAAT CGAAATAATC AAAATGAATA TGAAATTATT GATGCCCCCC	240
ATTGTGGGTG TCCATCAGAT GACGATGTGA GGTATCCTTT GGCAAGTGAC CCAAATGCAG	300
CGTTACAAAA TATGAACTAT AAAGATTACT TACAAATGAC AGATGAGGAC TACACTGATT	360
CTTATATAAA TCCTAGTTTA TCTATTAGTG GTAGAGATGC AGTTCAGACT GCGCTTACTG	420
TTGTTGGGAG AATACTCGGG GCTTTAGGTG TTCCGTTTTT TGGACAAATA GTGAGTTTTT	480
ATCAATTCCT TTAAATACA CTGTGGCCAG TTAATGATAC AGCTATATGG GAAGCTTTCA	540
TGCGACAGGT GGAGGAACTT GTCAATCAAC AAATAACAGA ATTTGCAAGA AATCAGGCAC	600
TTGCAAGATT GCAAGGATTA GGAGACTCTT TTAATGTATA TCAACGTTCC CTTCAAAATT	660
GGTTGGCTGA TCGAAATGAT ACACGAAATT TAAGTGTTGT TCGTGCTCAA TTTATAGCTT	720
TAGACCTTGA TTTTGTTAAT GCTATTCCAT TGTTTGCAGT AAATGGACAG CAGGTTCCAT	780
TACTGTCAGT ATATGCACAA GCTGTGAATT TACATTTGTT ATTATTAAAA GATGCATCTC	840
TTTTTGGA GAAGATGGGA TTCACACAGG GGGAAATTTT CACATATTAT GACCGTCAAT	900
TGGAACAAAC CGCTAAGTAC ACTAATTACT GTGAAACTTG GTATAATACA GGTTTAGATC	960
GTTTAAGAGG AACAAATACT GAAAGTTGGT TAAGATATCA TCAATTCCGT AGAGAAATGA	1020
CTTTAGTGGT ATTAGATGTT GTGGCGCTAT TTCCATATTA TGATGTACGA CTTTATCCAA	1080
CGGGATCAAA CCCACAGCTT ACACGTGAGG TATATACAGA TCCGATTGTA TTTAATCCAC	1140
CAGCTAATGT TGGACTTTGC CGACGTTGGG GTACTAATCC CTATAATACT TTTTCTGAGC	1200
TCGAAAATGC CTTCAATTCG CCACCACATC TTTTGTATAG GCTGAATAGC TTAACAATCA	1260

36

GCAGTAATCG	ATTTCCAGTT	TCATCTAATT	TTATGGATTA	TTGGTCAGGA	CATACGTTAC	1320
GCCGTAGTTA	TCTGAACGAT	TCAGCAGTAC	AAGAAGATAG	TTATGGCCTA	ATTACAACCA	1380
CAAGAGCAAC	AATTAATCCC	GGAGTTGATG	GAACAAACCG	CATAGAGTCA	ACGGCAGTAG	1440
ATTTTCGTTC	TGCATTGATA	GGTATATATG	GCGTGAATAG	AGCTTCTTTT	GTCCCAGGAG	1500
GCTTGTTTAA	TGGTACGACT	TCTCCTGCTA	ATGGAGGATG	TAGAGATCTC	TATGATACAA	1560
A						1561

## (4) INFORMATION FOR SEQ ID NO: 3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1554 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Bacillus thuringiensis
- (B) STRAIN: BTS02618A

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1146..1148
- (D) OTHER INFORMATION: /function= "Presumed translational stop codon of bTS02618A gene"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

AAAATTATCC AACATACATT TATCAAAAAG TAGATGCATC GGTGTTAAAG CCTTATACAC	60
GCTATAGACT AGATGGATTT GTGAAGNGTA GTCAAGATTT AGAAATTGAT CTCATCCACC	120
ATCATAAAGT CCATCTTGTA AAAAATGTAC CAGATAATTT AGTATCTGAT ACTTACTCAG	180
ATGGTTCTTG CAGCGGAATC AACCGTTGTG ATGAACAGCA TCAGGTAGAT ATGCAGCTAG	240
ATGCGGAGCA TCATCCAATG GATTGCTGTG AAGCGGCTCA AACACATGAG TTTTCTTCCT	300
ATATTAATAC AGGGGATCTA AATGCAAGTG TAGATCAGGG CATTTGGGTT GTATTAAAAG	360
TTCAACAAC AGATGGGTAT GCGACGTTAG GAAATCTTGA ATTGGTAGAG GTTGGGCCAT	420
TATCGGGTGA ATCTCTAGAA CGGGAACAAA GAGATAATGC GAAATGGAAT GCAGAGCTAG	480
GAAGAAAACG TGCAGAAATA GATCGTGTGT ATTTAGCTGC GAAACAAGCA ATTAATCATC	540
TGTTTGTAGA CTATCAAGAT CAACAATTAA ATCCAGAAAT TGGGCTAGCA GAAATTAATG	600
AAGCTTCAAA TCTTGTAGAG TCAATTTTCGG GTGTATATAG TGATACACTA TTACAGATTC	660
CTGGGATTAA CTACGAAATT TACACAGAGT TATCCGATCG CTTACAACAA GCATCGTATC	720
TGTATACGTC TAGAAATGCG GTGCAAAATG GAGACTTTAA CAGTGGTCTA GATAGTTGGA	780
ATACAACATAT GGATGCATCG GTTCAGCAAG ATGGCAATAT GCATTTCTTA GTTCTTTTCGC	840
ATTGGGATGC ACAAGTTTCC CAACAATTGA GAGTAAATCC GAATTGTAAG TATGTCTTAC	900
GTGTGACAGC AAGAAAAGTA GGAGGCGGAG ATGGATACGT CACAATCCGA GATGGCGCTC	960
ATCACCAAGA AACTCTTACA TTTAATGCAT GTGACTACGA TGTAATGGT ACGTATGTCA	1020
ATGACAATTC GTATATAACA GAAGAAGTGG TATTCTACCC AGAGACAAAA CATATGTGGG	1080
TAGAGGTGAG TGAATCCGAA GGTTCATTCT ATATAGACAG TATTGAGTTT ATTGAAACAC	1140
AAGAGTAGAA GAGGGGGATC CTAACGTATA GCAACTATGA GAGGATACTC CGTACAAACA	1200

AAGATTAAAA	AAAGGTAAAA	TGAATAGAAC	CCCCTACTGG	TAGAAGGACC	GATAGGGGGT	1260
TCTTACATGA	AAAAATGTAG	CTGTTTACTA	AGGTGTATAA	AAAACAGCAT	ATCTGATAGA	1320
AAAAAGTGAG	TACCTTATAA	AGAAAGAATT	CCATTCACAG	TTTCGGTATC	ATATAAATAA	1380
TGATAGGGGT	ATCCTTCTTA	TTTACATTAT	TTTTCGCAAT	TATCTCGACG	TTCTTCTTTC	1440
CGCTCACAAT	GATGATGATC	ATGACAACAA	TCGCGTCCAT	AGCGAACTCT	TTCGATATTA	1500
ATAATATCTA	AACTCGTGTA	GCAGTCATTT	CCATTTTTTT	TGATCCAGTA	AATA	1554

## (2) INFORMATION FOR SEQ ID NO: 4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4344 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

## (ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 668..4141

(D) OTHER INFORMATION: encompasses the entire sequence of SEQ ID NO 2: from nucleotide position 474 to 2034 in SEQ ID NO 4; also encompasses part of the sequence of SEQ ID NO 3: from nucleotide position 2994 to nucleotide position 4344 in SEQ ID NO 4; SEQ ID NO 3 shows additional nucleotides, located downstream (3') from the sequence shown in SEQ ID NO 4 (nucleotide position 1352 to nucleotide position 1554 in SEQ ID NO 3)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

GAATTCGAGC TCGGTACCTT TTCAGTGTAT CGTTTCCCTT CCATCAGGTT TTCAAATTGA      60
AAAGCCGAAT GATTTGAAAC TTGTTTACGA TGTAAGTCAT TTGTCTATGA CGAAAGATAC      120
GTGTAAAAAA CGTATTGAGA TTGATGAATG TGGACAAGTA GAAATTGACT TACAAGTATT      180
AAAGATTAAG GGTGTCCTTT CTTTATCGG AAATTTCTCT ATTGAACCTA TTCTGTGTGA      240
AAACATGTAT ACAACGGTTG ATAGAGATCC GTCTATTTCC TTAAGTTTCC AAGATACGGT      300
ATATGTGGAC CATATTTTAA AATATAGCGT CCAACAATA CCATATTATG TAATTGATGG      360
TGATCATATT CAAGTACGTG ATTTACAAAT CAACTGATG AAAGAGAATC CGCAATCTGC      420
TCAAGTATCA GGTGTGTTTT GTTTTGTATA TGAGTAAGAA CCGAAGGTTT GTAAAAAAGA      480
AATAGGAATA AATACTATCC ATTTTTCAT GAAATATTTT TTTATTAGAA AGGAATCTTT      540
CTTACACGGG AAAATCCTAA GATTGAGAGT AAAGATATAT ATATATAAAT ACAATAAAGA      600
GTTTGTGAGG ATTTTGTAAA GATATGATAT GAACATGCAC TAGATTTATA GTATAGGAGG      660
AAAAAGT ATG AAT CGA AAT AAT CAA AAT GAA TAT GAA ATT ATT GAT GCC      709
      Met Asn Arg Asn Asn Gln Asn Glu Tyr Glu Ile Ile Asp Ala
          1              5              10

CCC CAT TGT GGG TGT CCA TCA GAT GAC GAT GTG AGG TAT CCT TTG GCA      757
Pro His Cys Gly Cys Pro Ser Asp Asp Asp Val Arg Tyr Pro Leu Ala
    15              20              25              30

AGT GAC CCA AAT GCA GCG TTA CAA AAT ATG AAC TAT AAA GAT TAC TTA      805
Ser Asp Pro Asn Ala Ala Leu Gln Asn Met Asn Tyr Lys Asp Tyr Leu
          35              40              45

CAA ATG ACA GAT GAG GAC TAC ACT GAT TCT TAT ATA AAT CCT AGT TTA      853
Gln Met Thr Asp Glu Asp Tyr Thr Asp Ser Tyr Ile Asn Pro Ser Leu
          50              55              60

```

40

TCT	ATT	AGT	GGT	AGA	GAT	GCA	GTT	CAG	ACT	GCG	CTT	ACT	GTT	GTT	GGG	901
Ser	Ile	Ser	Gly	Arg	Asp	Ala	Val	Gln	Thr	Ala	Leu	Thr	Val	Val	Gly	
		65					70					75				
AGA	ATA	CTC	GGG	GCT	TTA	GGT	GTT	CCG	TTT	TCT	GGA	CAA	ATA	GTG	AGT	949
Arg	Ile	Leu	Gly	Ala	Leu	Gly	Val	Pro	Phe	Ser	Gly	Gln	Ile	Val	Ser	
		80				85					90					
TTT	TAT	CAA	TTC	CTT	TTA	AAT	ACA	CTG	TGG	CCA	GTT	AAT	GAT	ACA	GCT	997
Phe	Tyr	Gln	Phe	Leu	Leu	Asn	Thr	Leu	Trp	Pro	Val	Asn	Asp	Thr	Ala	
95					100					105					110	
ATA	TGG	GAA	GCT	TTC	ATG	CGA	CAG	GTG	GAG	GAA	CTT	GTC	AAT	CAA	CAA	1045
Ile	Trp	Glu	Ala	Phe	Met	Arg	Gln	Val	Glu	Glu	Leu	Val	Asn	Gln	Gln	
				115					120					125		
ATA	ACA	GAA	TTT	GCA	AGA	AAT	CAG	GCA	CTT	GCA	AGA	TTG	CAA	GGA	TTA	1093
Ile	Thr	Glu	Phe	Ala	Arg	Asn	Gln	Ala	Leu	Ala	Arg	Leu	Gln	Gly	Leu	
			130				135						140			
GGA	GAC	TCT	TTT	AAT	GTA	TAT	CAA	CGT	TCC	CTT	CAA	AAT	TGG	TTG	GCT	1141
Gly	Asp	Ser	Phe	Asn	Val	Tyr	Gln	Arg	Ser	Leu	Gln	Asn	Trp	Leu	Ala	
		145					150					155				
GAT	CGA	AAT	GAT	ACA	CGA	AAT	TTA	AGT	GTT	GTT	CGT	GCT	CAA	TTT	ATA	1189
Asp	Arg	Asn	Asp	Thr	Arg	Asn	Leu	Ser	Val	Val	Arg	Ala	Gln	Phe	Ile	
	160					165					170					
GCT	TTA	GAC	CTT	GAT	TTT	GTT	AAT	GCT	ATT	CCA	TTG	TTT	GCA	GTA	AAT	1237
Ala	Leu	Asp	Leu	Asp	Phe	Val	Asn	Ala	Ile	Pro	Leu	Phe	Ala	Val	Asn	
175					180					185					190	
GGA	CAG	CAG	GTT	CCA	TTA	CTG	TCA	GTA	TAT	GCA	CAA	GCT	GTG	AAT	TTA	1285
Gly	Gln	Gln	Val	Pro	Leu	Leu	Ser	Val	Tyr	Ala	Gln	Ala	Val	Asn	Leu	
				195					200					205		
CAT	TTG	TTA	TTA	TTA	AAA	GAT	GCA	TCT	CTT	TTT	GGA	GAA	GGA	TGG	GGA	1333
His	Leu	Leu	Leu	Leu	Lys	Asp	Ala	Ser	Leu	Phe	Gly	Glu	Gly	Trp	Gly	
			210					215					220			
TTC	ACA	CAG	GGG	GAA	ATT	TCC	ACA	TAT	TAT	GAC	CGT	CAA	TTG	GAA	CTA	1381
Phe	Thr	Gln	Gly	Glu	Ile	Ser	Thr	Tyr	Tyr	Asp	Arg	Gln	Leu	Glu	Leu	
		225					230					235				
ACC	GCT	AAG	TAC	ACT	AAT	TAC	TGT	GAA	ACT	TGG	TAT	AAT	ACA	GGT	TTA	1429
Thr	Ala	Lys	Tyr	Thr	Asn	Tyr	Cys	Glu	Thr	Trp	Tyr	Asn	Thr	Gly	Leu	
	240					245					250					
GAT	CGT	TTA	AGA	GGA	ACA	AAT	ACT	GAA	AGT	TGG	TTA	AGA	TAT	CAT	CAA	1477
Asp	Arg	Leu	Arg	Gly	Thr	Asn	Thr	Glu	Ser	Trp	Leu	Arg	Tyr	His	Gln	
255					260					265					270	
TTC	CGT	AGA	GAA	ATG	ACT	TTA	GTG	GTA	TTA	GAT	GTT	GTG	GCG	CTA	TTT	1525
Phe	Arg	Arg	Glu	Met	Thr	Leu	Val	Val	Leu	Asp	Val	Val	Ala	Leu	Phe	
				275					280					285		
CCA	TAT	TAT	GAT	GTA	CGA	CTT	TAT	CCA	ACG	GGA	TCA	AAC	CCA	CAG	CTT	1573
Pro	Tyr	Tyr	Asp	Val	Arg	Leu	Tyr	Pro	Thr	Gly	Ser	Asn	Pro	Gln	Leu	
			290					295					300			

ACA CGT GAG GTA TAT ACA GAT CCG ATT GTA TTT AAT CCA CCA GCT AAT	1621
Thr Arg Glu Val Tyr Thr Asp Pro Ile Val Phe Asn Pro Pro Ala Asn	
305 310 315	
GTT GGA CTT TGC CGA CGT TGG GGT ACT AAT CCC TAT AAT ACT TTT TCT	1669
Val Gly Leu Cys Arg Arg Trp Gly Thr Asn Pro Tyr Asn Thr Phe Ser	
320 325 330	
GAG CTC GAA AAT GCC TTC ATT CGC CCA CCA CAT CTT TTT GAT AGG CTG	1717
Glu Leu Glu Asn Ala Phe Ile Arg Pro Pro His Leu Phe Asp Arg Leu	
335 340 345 350	
AAT AGC TTA ACA ATC AGC AGT AAT CGA TTT CCA GTT TCA TCT AAT TTT	1765
Asn Ser Leu Thr Ile Ser Ser Asn Arg Phe Pro Val Ser Ser Asn Phe	
355 360 365	
ATG GAT TAT TGG TCA GGA CAT ACG TTA CGC CGT AGT TAT CTG AAC GAT	1813
Met Asp Tyr Trp Ser Gly His Thr Leu Arg Arg Ser Tyr Leu Asn Asp	
370 375 380	
TCA GCA GTA CAA GAA GAT AGT TAT GGC CTA ATT ACA ACC ACA AGA GCA	1861
Ser Ala Val Gln Glu Asp Ser Tyr Gly Leu Ile Thr Thr Thr Arg Ala	
385 390 395	
ACA ATT AAT CCC GGA GTT GAT GGA ACA AAC CGC ATA GAG TCA ACG GCA	1909
Thr Ile Asn Pro Gly Val Asp Gly Thr Asn Arg Ile Glu Ser Thr Ala	
400 405 410	
GTA GAT TTT CGT TCT GCA TTG ATA GGT ATA TAT GGC GTG AAT AGA GCT	1957
Val Asp Phe Arg Ser Ala Leu Ile Gly Ile Tyr Gly Val Asn Arg Ala	
415 420 425 430	
TCT TTT GTC CCA GGA GGC TTG TTT AAT GGT ACG ACT TCT CCT GCT AAT	2005
Ser Phe Val Pro Gly Gly Leu Phe Asn Gly Thr Thr Ser Pro Ala Asn	
435 440 445	
GGA GGA TGT AGA GAT CTC TAT GAT ACA AAT GAT GAA TTA CCA CCA GAT	2053
Gly Gly Cys Arg Asp Leu Tyr Asp Thr Asn Asp Glu Leu Pro Pro Asp	
450 455 460	
GAA AGT ACC GGA AGT TCA ACC CAT AGA CTA TCT CAT GTT ACC TTT TTT	2101
Glu Ser Thr Gly Ser Ser Thr His Arg Leu Ser His Val Thr Phe Phe	
465 470 475	
AGC TTT CAA ACT AAT CAG GCT GGA TCT ATA GCT AAT GCA GGA AGT GTA	2149
Ser Phe Gln Thr Asn Gln Ala Gly Ser Ile Ala Asn Ala Gly Ser Val	
480 485 490	
CCT ACT TAT GTT TGG ACC CGT CGT GAT GTG GAC CTT AAT AAT ACG ATT	2197
Pro Thr Tyr Val Trp Thr Arg Arg Asp Val Asp Leu Asn Asn Thr Ile	
495 500 505 510	
ACC CCA AAT AGA ATT ACA CAA TTA CCA TTG GTA AAG GCA TCT GCA CCT	2245
Thr Pro Asn Arg Ile Thr Gln Leu Pro Leu Val Lys Ala Ser Ala Pro	
515 520 525	
GTT TCG GGT ACT ACG GTC TTA AAA GGT CCA GGA TTT ACA GGA GGG GGT	2293
Val Ser Gly Thr Val Leu Lys Gly Pro Gly Phe Thr Gly Gly Gly	
530 535 540	



ATA CTC CGA AGA ACA ACT AAT GGC ACA TTT GGA ACG TTA AGA GTA ACG Ile Leu Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu Arg Val Thr 545 550 555	2341
GTT AAT TCA CCA TTA ACA CAA CAA TAT CGC CTA AGA GTT CGT TTT GCC Val Asn Ser Pro Leu Thr Gln Gln Tyr Arg Leu Arg Val Arg Phe Ala 560 565 570	2389
TCA ACA GGA AAT TTC AGT ATA AGG GTA CTC CGT GGA GGG GTT TCT ATC Ser Thr Gly Asn Phe Ser Ile Arg Val Leu Arg Gly Gly Val Ser Ile 575 580 585 590	2437
GGT GAT GTT AGA TTA GGG AGC ACA ATG AAC AGA GGG CAG GAA CTA ACT Gly Asp Val Arg Leu Gly Ser Thr Met Asn Arg Gly Gln Glu Leu Thr 595 600 605	2485
TAC GAA TCC TTT TTC ACA AGA GAG TTT ACT ACT ACT GGT CCG TTC AAT Tyr Glu Ser Phe Phe Thr Arg Glu Phe Thr Thr Thr Gly Pro Phe Asn 610 615 620	2533
CCG CCT TTT ACA TTT ACA CAA GCT CAA GAG ATT CTA ACA GTG AAT GCA Pro Pro Phe Thr Phe Thr Gln Ala Gln Glu Ile Leu Thr Val Asn Ala 625 630 635	2581
GAA GGT GTT AGC ACC GGT GGT GAA TAT TAT ATA GAT AGA ATT GAA ATT Glu Gly Val Ser Thr Gly Gly Glu Tyr Tyr Ile Asp Arg Ile Glu Ile 640 645 650	2629
GTC CCT GTG AAT CCG GCA CGA GAA GCG GAA GAG GAT TTA GAA GCG GCG Val Pro Val Asn Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Ala Ala 655 660 665 670	2677
AAG AAA GCG GTG GCG AGC TTG TTT ACA CGT ACA AGG GAC GGA TTA CAG Lys Lys Ala Val Ala Ser Leu Phe Thr Arg Thr Arg Asp Gly Leu Gln 675 680 685	2725
GTA AAT GTG ACA GAT TAT CAA GTG GAC CAA GCG GCA AAT TTA GTG TCA Val Asn Val Thr Asp Tyr Gln Val Asp Gln Ala Ala Asn Leu Val Ser 690 695 700	2773
TGC TTA TCC GAT GAA CAA TAT GGG CAT GAC AAA AAG ATG TTA TTG GAA Cys Leu Ser Asp Glu Gln Tyr Gly His Asp Lys Lys Met Leu Leu Glu 705 710 715	2821
GCG GTA AGA GCG GCA AAA CGC CTC AGC CGC GAA CGC AAC TTA CTT CAA Ala Val Arg Ala Ala Lys Arg Leu Ser Arg Glu Arg Asn Leu Leu Gln 720 725 730	2869
GAT CCA GAT TTT AAT ACA ATC AAT AGT ACA GAA GAG AAT GGC TGG AAG Asp Pro Asp Phe Asn Thr Ile Asn Ser Thr Glu Glu Asn Gly Trp Lys 735 740 745 750	2917
GCA AGT AAC GGT GTT ACT ATT AGC GAG GGC GGT CCA TTC TTT AAA GGT Ala Ser Asn Gly Val Thr Ile Ser Glu Gly Gly Pro Phe Phe Lys Gly 755 760 765	2965
CGT GCA CTT CAG TTA GCA AGC GCA AGA GAA AAT TAT CCA ACA TAC ATT Arg Ala Leu Gln Leu Ala Ser Ala Arg Glu Asn Tyr Pro Thr Tyr Ile 770 775 780	3013

TAT CAA AAA GTA GAT GCA TCG GTG TTA AAG CCT TAT ACA CGC TAT AGA Tyr Gln Lys Val Asp Ala Ser Val Leu Lys Pro Tyr Thr Arg Tyr Arg 785 790 795	3061
CTA GAT GGA TTT GTG AAG AGT AGT CAA GAT TTA GAA ATT GAT CTC ATC Leu Asp Gly Phe Val Lys Ser Ser Gln Asp Leu Glu Ile Asp Leu Ile 800 805 810	3109
CAC CAT CAT AAA GTC CAT CTT GTA AAA AAT GTA CCA GAT AAT TTA GTA His His His Lys Val His Leu Val Lys Asn Val Pro Asp Asn Leu Val 815 820 825 830	3157
TCT GAT ACT TAC TCA GAT GGT TCT TGC AGC GGA ATC AAC CGT TGT GAT Ser Asp Thr Tyr Ser Asp Gly Ser Cys Ser Gly Ile Asn Arg Cys Asp 835 840 845	3205
GAA CAG CAT CAG GTA GAT ATG CAG CTA GAT GCG GAG CAT CAT CCA ATG Glu Gln His Gln Val Asp Met Gln Leu Asp Ala Glu His His Pro Met 850 855 860	3253
GAT TGC TGT GAA GCG GCT CAA ACA CAT GAG TTT TCT TCC TAT ATT AAT Asp Cys Cys Glu Ala Ala Gln Thr His Glu Phe Ser Ser Tyr Ile Asn 865 870 875	3301
ACA GGG GAT CTA AAT GCA AGT GTA GAT CAG GGC ATT TGG GTT GTA TTA Thr Gly Asp Leu Asn Ala Ser Val Asp Gln Gly Ile Trp Val Val Leu 880 885 890	3349
AAA GTT CGA ACA ACA GAT GGG TAT GCG ACG TTA GGA AAT CTT GAA TTG Lys Val Arg Thr Thr Asp Gly Tyr Ala Thr Leu Gly Asn Leu Glu Leu 895 900 905 910	3397
GTA GAG GTT GGG CCA TTA TCG GGT GAA TCT CTA GAA CGG GAA CAA AGA Val Glu Val Gly Pro Leu Ser Gly Glu Ser Leu Glu Arg Glu Gln Arg 915 920 925	3445
GAT AAT GCG AAA TGG AAT GCA GAG CTA GGA AGA AAA CGT GCA GAA ATA Asp Asn Ala Lys Trp Asn Ala Glu Leu Gly Arg Lys Arg Ala Glu Ile 930 935 940	3493
GAT CGT GTG TAT TTA GCT GCG AAA CAA GCA ATT AAT CAT CTG TTT GTA Asp Arg Val Tyr Leu Ala Ala Lys Gln Ala Ile Asn His Leu Phe Val 945 950 955	3541
GAC TAT CAA GAT CAA CAA TTA AAT CCA GAA ATT GGG CTA GCA GAA ATT Asp Tyr Gln Asp Gln Gln Leu Asn Pro Glu Ile Gly Leu Ala Glu Ile 960 965 970	3589
AAT GAA GCT TCA AAT CTT GTA GAG TCA ATT TCG GGT GTA TAT AGT GAT Asn Glu Ala Ser Asn Leu Val Glu Ser Ile Ser Gly Val Tyr Ser Asp 975 980 985 990	3637
ACA CTA TTA CAG ATT CCT GGG ATT AAC TAC GAA ATT TAC ACA GAG TTA Thr Leu Leu Gln Ile Pro Gly Ile Asn Tyr Glu Ile Tyr Thr Thr Glu Leu 995 1000 1005	3685
TCC GAT CGC TTA CAA CAA GCA TCG TAT CTG TAT ACG TCT AGA AAT GCG Ser Asp Arg Leu Gln Gln Ala Ser Tyr Leu Tyr Thr Ser Arg Asn Ala 1010 1015 1020	3733

GTG CAA AAT GGA GAC TTT AAC AGT GGT CTA GAT AGT TGG AAT ACA ACT Val Gln Asn Gly Asp Phe Asn Ser Gly Leu Asp Ser Trp Asn Thr Thr 1025 1030 1035	3781
ATG GAT GCA TCG GTT CAG CAA GAT GGC AAT ATG CAT TTC TTA GTT CTT Met Asp Ala Ser Val Gln Gln Asp Gly Asn Met His Phe Leu Val Leu 1040 1045 1050	3829
TCG CAT TGG GAT GCA CAA GTT TCC CAA CAA TTG AGA GTA AAT CCG AAT Ser His Trp Asp Ala Gln Val Ser Gln Gln Leu Arg Val Asn Pro Asn 1055 1060 1065 1070	3877
TGT AAG TAT GTC TTA CGT GTG ACA GCA AGA AAA GTA GGA GGC GGA GAT Cys Lys Tyr Val Leu Arg Val Thr Ala Arg Lys Val Gly Gly Gly Asp 1075 1080 1085	3925
GGA TAC GTC ACA ATC CGA GAT GGC GCT CAT CAC CAA GAA ACT CTT ACA Gly Tyr Val Thr Ile Arg Asp Gly Ala His His Gln Glu Thr Leu Thr 1090 1095 1100	3973
TTT AAT GCA TGT GAC TAC GAT GTA AAT GGT ACG TAT GTC AAT GAC AAT Phe Asn Ala Cys Asp Tyr Asp Val Asn Gly Thr Tyr Val Asn Asp Asn 1105 1110 1115	4021
TCG TAT ATA ACA GAA GAA GTG GTA TTC TAC CCA GAG ACA AAA CAT ATG Ser Tyr Ile Thr Glu Glu Val Val Phe Tyr Pro Glu Thr Lys His Met 1120 1125 1130	4069
TGG GTA GAG GTG AGT GAA TCC GAA GGT TCA TTC TAT ATA GAC AGT ATT Trp Val Glu Val Ser Glu Ser Glu Gly Ser Phe Tyr Ile Asp Ser Ile 1135 1140 1145 1150	4117
GAG TTT ATT GAA ACA CAA GAG TAGAAGAGGG GGATCCTAAC GTATAGCAAC Glu Phe Ile Glu Thr Gln Glu 1155	4168
TATGAGAGGA TACTCCGTAC AAACAAAGAT TAAAAAAGG TAAAATGAAT AGAACCCCCT	4228
ACTGGTAGAA GGACCGATAG GGGGTTCTTA CATGAAAAAA TGTAAGCTGTT TACTAAGGTG	4288
TATAAAAAAC AGCATATCTG ATAGAAAAAA GTGAGTACCT TATAAAGAAA GAATTC	4344